

Apparatus for adapting the directional characteristic of microphones for voice control

Patent number: DE19943872
Publication date: 2001-03-15
Inventor: SCHROEDER ERNST F (DE)
Applicant: THOMSON BRANDT GMBH (DE)
Classification:
- **international:** G10L15/00
- **european:** G10L15/20, H04R3/00B
Application number: DE19991043872 19990914
Priority number(s): DE19991043872 19990914

Also published as:

EP1085781 (A2)
JP2001125594 (A)
EP1085781 (A3)

Abstract not available for DE19943872

Abstract of correspondent: **EP1085781**

Voice control systems are used in diverse technical fields. In this case, the spoken words are firstly detected, usually by one or more microphones (M1, M2) and then fed to a speech recognition system (REC), which determines from the detected signals those words having the highest probability. The robustness with respect to acoustic interfering influences can be increased by the microphone being fastened or held directly in front of the speaker's mouth. The invention enables voice control even from a relatively large distance by virtue of the fact that interfering background noises are spatially separated by the use of directional microphones or microphone arrays (M1, M2). In this case, the directional characteristic is adapted by virtue of the fact that means for detecting the whereabouts of the user (IR1, VR, CIR) are provided. One or more infrared motion detectors are preferably used for this purpose.

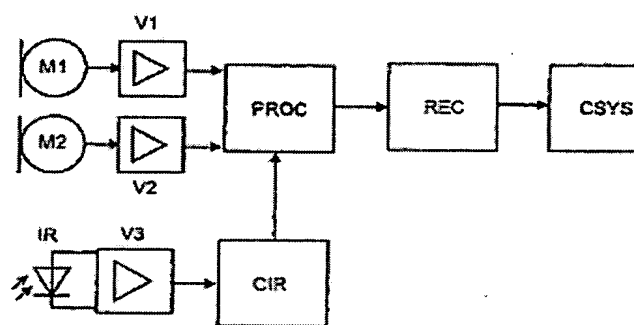


Fig. 1

Data supplied from the **esp@cenet** database - Worldwide

THIS PAGE BLANK (USPTO)



19 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENT- UND
MARKENAMT

12 Offenlegungsschrift
10 DE 199 43 872 A 1

57 Int. Cl. 7:
G 10 L 15/00

21 Aktenzeichen: 199 43 872.2
22 Anmeldetag: 14. 9. 1999
43 Offenlegungstag: 15. 3. 2001

DE 199 43 872 A 1

71 Anmelder:
Deutsche Thomson-Brandt GmbH, 78048
Villingen-Schwenningen, DE

72 Erfinder:
Schröder, Ernst F., Dr., 30655 Hannover, DE

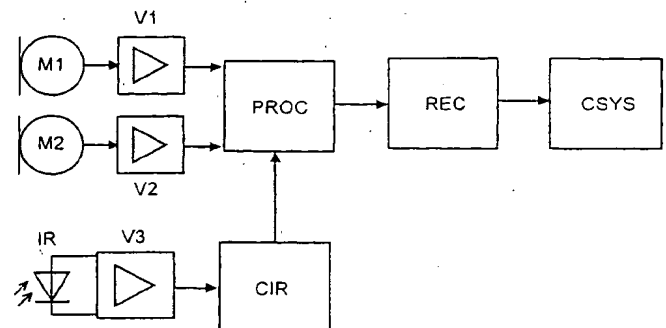
56 Für die Beurteilung der Patentfähigkeit in Betracht
zu ziehende Druckschriften:

DE 198 44 748 A1
DE 198 12 697 A1
DE 197 41 596 A1
DE 197 12 632 A1
DE 44 39 146 A1
DE 43 13 256 A1
DE 691 01 527 T2

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

54 Vorrichtung zur Anpassung der Richtcharakteristik von Mikrofonen für die Sprachsteuerung

57 Sprachsteuerungssysteme finden in einer Vielfalt von technischen Gebieten Anwendung. Die gesprochenen Worte werden hierbei zunächst detektiert, üblicherweise durch ein oder mehrere Mikrofone (M1, M2), und dann einem Spracherkennungssystem (REC) zugeführt, welches aus den detektierten Signalen diejenigen Worte bestimmt, welche die höchste Wahrscheinlichkeit aufweisen. Die Robustheit gegen akustische Störeinflüsse kann erhöht werden, indem das Mikrophon direkt vor dem Mund des Sprechers befestigt oder gehalten wird. Erfindungsgemäß wird eine Sprachsteuerung auch aus größerer Ferne ermöglicht, indem eine räumliche Trennung von störenden Hintergrundgeräuschen durch eine Verwendung von Richtmikrofonen bzw. Mikrofonarrays (M1, M2) erfolgt. Eine Anpassung der Richtcharakteristik erfolgt hierbei, indem Mittel zur Detektion des Aufenthaltsortes des Benutzers (IR1, VR, CIR) vorgesehen sind. Vorzugsweise werden hierfür ein oder mehrere Infrarot-Bewegungsdetektoren verwendet.



DE 199 43 872 A 1

Beschreibung

Die Erfindung betrifft eine Vorrichtung zur Anpassung der Richtcharakteristik von Mikrofonen für die Sprachsteuerung, die insbesondere zur Steuerung von Geräten der Unterhaltungselektronik genutzt werden kann.

Stand der Technik

Sprachsteuerungssysteme finden in einer Vielzahl von technischen Gebieten Anwendung. Die gesprochenen Worte werden hierbei zunächst als Schallsignale detektiert, üblicherweise durch ein oder mehrere Mikrofone, und dann einem Spracherkennungssystem zugeführt. Die Spracherkennung basiert hierbei üblicherweise auf einem Akustik- und einem Sprachmodell. Das akustische Modell nutzt eine große Anzahl von Sprachmustern, wobei mathematische Algorithmen dazu verwendet werden, die akustisch am besten passenden Worte zu einem gesprochenen Wort anzugeben. Das Sprachmodell wiederum basiert auf einer Analyse, bei der anhand von einer Vielzahl von Dokumentproben festgestellt wird, in welchem Kontext und wie häufig gewisse Wörter normalerweise verwendet werden. Mit solchen Spracherkennungssystemen ist nicht nur das Erkennen einzelner Wörter, sondern auch von fließend gesprochenen Sätzen mit hohen Erkennungsraten möglich. Die Erkennungsrate sinkt jedoch drastisch, wenn nicht vernachlässigbare Hintergrundgeräusche vorliegen.

Die Robustheit gegen solche akustische Störeinflüsse kann auf verschiedene Weisen erhöht werden. So wird bei Diktiersystemen für Computer ein Mikrofon an einem Kopfhörergestell direkt vor dem Mund des Sprechers befestigt. Bei diesen Systemen kann nur durch die unmittelbare Nähe zum Mund ein sehr konstantes Signal und damit eine zum Teil beachtliche Erkennungsrate erreicht werden. Ebenso ist es bekannt, ein Fernsehgerät zu steuern, indem die Bedienungsbefehle in das in einer Fernbedienung integrierte Mikrofon gesprochen werden. Auch hier muß jedoch die Fernbedienung unmittelbar vor den Mund des Benutzers gehalten werden.

Erfindung

Der Erfindung liegt die Aufgabe zugrunde, eine Vorrichtung zur Sprachsteuerung anzugeben, welche eine ausreichende Störfestigkeit auch bei Spracheingabe aus größerer Ferne ermöglicht. Diese Aufgabe wird durch die in Anspruch 1 angegebene Vorrichtung gelöst.

Um eine Sprachsteuerung auch aus größerer Ferne zu ermöglichen, muß das Sprachsignal von störenden Hintergrundsignalen getrennt werden. Dies kann durch eine räumliche Trennung unter Verwendung von Richtmikrofonen bzw. Mikrofonarrays aus zwei oder mehr Mikrofonen erfolgen. Soll der Benutzer sich während der Spracheingabe jedoch frei im Raum bewegen können statt auf einen begrenzten Bereich, z. B. mittig vor dem Gerät, beschränkt zu sein, so muß eine Anpassung der Richtcharakteristik dieser Mikrofonarrays erfolgen. Für diese Anpassung könnten die Mikrofonsignale selber verwendet werden. In diesem Fall muß zunächst detektiert werden, daß ein entsprechendes Signal vorliegt, dann kann das Mikrofonarray entsprechend adaptiert werden, und erst im Anschluß kann eine zufriedenstellende Spracherkennung beginnen. Dieser Verfahrensablauf kann jedoch mehrere Sekunden benötigen und führt damit dazu, daß die ersten Silben oder sogar Wörter nicht erkannt werden können.

Im Prinzip besteht die erfindungsgemäße Vorrichtung zur Sprachfernsteuerung aus einem oder mehreren Mikrofonen

zur Umwandlung von Sprachbefehlen in elektrische Signale und einer Spracherkennungseinheit zur Umwandlung dieser elektrischen Signale in Bedienungsbefehle sowie Mitteln zur Detektion des Aufenthaltsortes des Benutzers, die ein Signal zur Verfügung stellen, welches zur Anpassung der Richtcharakteristik des Mikrofons oder der Mikrofone an den Aufenthaltsort des Benutzers verwendet werden kann.

Die Verwendung eines separaten Detektionsmittels hat den Vorteil, daß eine schnelle Erfassung des Aufenthaltsortes des Benutzers stets auch ein Detektieren und Erkennen der ersten Silben bzw. Wörter eines Kommandos ermöglicht. Hierfür ist grundsätzlich jedes ausreichend schnelle Detektionsmittel geeignet.

Vorteilhaft weisen die Mittel zur Detektion des Aufenthaltsortes des Benutzers ein oder mehrere Infrarot-Bewegungsdetektoren auf.

Ebenso kann es vorteilhaft sein, wenn die Mittel zur Detektion des Aufenthaltsortes des Benutzers zusätzlich zu den Infrarot-Bewegungsdetektoren oder statt der Infrarot-Bewegungsdetektoren elektrische Kontakte aufweisen.

Vorzugsweise kann auch ein Richtmikrofon vorgesehen sein, welches auf den Aufenthaltsort des Benutzers ausgerichtet wird.

Besonders vorteilhaft ist die erfindungsgemäße Vorrichtung in ein Unterhaltungselektronikgerät integriert.

Zeichnung

Anhand der Zeichnung wird ein Ausführungsbeispiel der Erfindung beschrieben.

Diese zeigt eine erfindungsgemäße Anordnung zur Sprachsteuerung.

Ausführungsbeispiele

In Fig. 1 ist schematisch eine erfindungsgemäße Vorrichtung dargestellt, wie sie z. B. in einem Fernsehgerät, Videorekorder oder DVD-Spieler integriert werden kann. Um die Position eines Benutzers zu detektieren, bevor dieser einen Benutzungsbefehl gegeben hat, ist ein Infrarot-Positionsdetektor vorgesehen. Dieser beruht darauf, daß die Intensität der von einem Gegenstand oder einer Person abgestrahlten Infrarotstrahlung von der jeweiligen Temperatur abhängt. Der Infrarot-Positionsdetektor umfaßt eine Vielzahl von dem Fachmann bekannten und daher nicht näher beschriebenen Infrarotdetektoren IR, da mit einem einzelnen Detektor lediglich die Anwesenheit oder Bewegung eines Benutzers, nicht jedoch eine Positionsbestimmung erfolgen kann. Je nach Anforderungen an die Auflösung kann es sich um mehrere einzelne Detektoren oder ein Array, wie es auch für Infrarot- bzw. Wärmebildkameras benutzt wird, handeln. Durch eine nicht dargestellte Linse oder Linsenkombination kann hierbei der Beobachtungsbereich beeinflußt werden. Die aus der Infrarotstrahlung durch die Infrarotdetektoren gewonnenen elektrischen Signale werden dann durch einen Verstärker V3 verstärkt und gelangen zu einer Kontrolleinheit CIR, die aus den detektierten Signalen den Aufenthaltsort und gegebenenfalls auch die Bewegungsrichtung des Benutzers ermittelt.

Zur Detektion der Sprachsignale ist ein Mikrofonarray bestehend aus zwei Mikrofonen M1 und M2 vorgesehen. Diese setzen die detektierten Schallsignale in elektrische Signale um, die durch Verstärker V1 und V2 zunächst verstärkt und dann einer Signalverarbeitungseinheit PROC zugeführt werden. Diese berücksichtigt nun den jeweiligen Aufenthaltsort des Benutzers durch eine unterschiedliche Skalierung bzw. Verarbeitung der detektierten Schallsignale. Das so bearbeitete Signal wird dann einer Spracher-

kennungseinheit REC zugeführt, die die elektrischen Signale in Worte umwandelt. Die diesen Worten entsprechenden Befehle werden dann schließlich einem Systemmanager CSYS zur Steuerung des Systems zugeführt.

Bei einem aus zwei Mikrofonen bestehenden Mikrofonarray können die beiden Mikrofone vorteilhaft auf der linken und rechten Seite des Gehäuses des jeweiligen Gerätes untergebracht werden. Das Mikrofonarray ist jedoch keinesfalls auf zwei Mikrofone beschränkt, sondern kann auch mehr als zwei Mikrofone aufweisen. Ebenso kann auch ein Richtmikrofon mit einer ausgeprägten Richtcharakteristik verwendet werden, welches dann in Abhängigkeit von dem detektierten Aufenthaltsort des Benutzers verdreht wird.

Vorteilhafterweise kann in einer Grundeinstellung eine Ausrichtung in Richtung des Eingangsbereiches des Raumes, in dem sich das Gerät befindet, erfolgen. Damit ist das Gerät sofort einsatzbereit, sobald ein Benutzer den Raum betritt.

Sollten mehrere mögliche Benutzer detektiert werden, so sind verschiedene Verfahrensweisen denkbar. Beispielsweise kann die Richtcharakteristik stets an den zuerst detektierten Benutzer angepaßt werden, so daß das Mikrofonarray diesem gewissermaßen folgt. Ebenso kann das Mikrofonarray auch demjenigen Benutzer folgen, der zuerst eine Spracheingabe vorgenommen hat. Weiterhin kann auch ein regelmäßiges Umschalten der Ausrichtung zwischen den verschiedenen möglichen Benutzern erfolgen. Schließlich kann eine Auswertung der Größe der detektierten Objekte erfolgen, um so zu verhindern, daß eine Ausrichtung auf Haustiere des Benutzers durchgeführt wird.

Weiterhin ist es denkbar, den Benutzer beim Betreten des Raumes statt durch den Infrarotdetektor durch Kontakte an der Tür oder an den Lichtschaltern zu detektieren.

Schließlich ist auch eine Kombination solcher Kontakte mit einer Infrarotdetektion möglich.

Gemäß einer weiteren Ausführungsform können bei ausgeschalteten Geräten die Detektionsmittel ständig aktiviert sein. Erfolgt dann eine Detektion eines Benutzers, so wird zunächst lediglich die Richtcharakteristik der Mikrofone angepaßt und die Spracherkennungseinheit angeschaltet. Erfolgt dann ein Benutzungsbefehl, insbesondere ein Kommando zum Anschalten des zu steuernden Geräts, so werden auch die restlichen Funktionsgruppen des Gerätes aktiviert. Auf diese Weise ist das Gerät ständig benutzungsbereit bei lediglich geringem Energiebedarf des Geräts.

Die Erfindung kann zur Sprachfernbedienung von verschiedensten Geräten der Unterhaltungselektronik, wie z. B. von TV-Geräten, Videorecordern, DVD-Spielern, Satellitenempfängern, TV-Video-Kombinationen, Audiogeräten oder kompletten Audiosystemen, aber ebenso von Personalcomputern oder von Haushaltsgeräten eingesetzt werden.

3. Vorrichtung nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß die Mittel zur Detektion des Aufenthaltsortes des Benutzers elektrische Kontakte aufweisen.

4. Vorrichtung nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß ein Richtmikrofon vorgesehen ist, welches auf den Aufenthaltsort des Benutzers ausgerichtet wird.

5. Unterhaltungselektronikgerät mit einer Vorrichtung nach einem der vorhergehenden Ansprüche.

Hierzu 1 Seite(n) Zeichnungen

Patentansprüche

1. Vorrichtung zur Sprachfernsteuerung, mit einem oder mehreren Mikrofonen (M1, M2) zur Umwandlung von Sprachbefehlen in elektrische Signale und mit einer Spracherkennungseinheit (REC) zur Umwandlung dieser elektrischen Signale in Bedienungsbefehle, **dadurch gekennzeichnet**, daß Mittel zur Detektion des Aufenthaltsortes des Benutzers (IR, V3, CIR) vorgesehen sind, die ein Signal zur Verfügung stellen, welches zur Anpassung der Richtcharakteristik des Mikrofons oder der Mikrofone an den Aufenthaltsort des Benutzers verwendet werden kann.

2. Vorrichtung nach Anspruch 1, dadurch gekennzeichnet, daß die Mittel zur Detektion des Aufenthaltsortes des Benutzers ein oder mehrere Infrarot-Bewegungsdetektoren aufweisen.

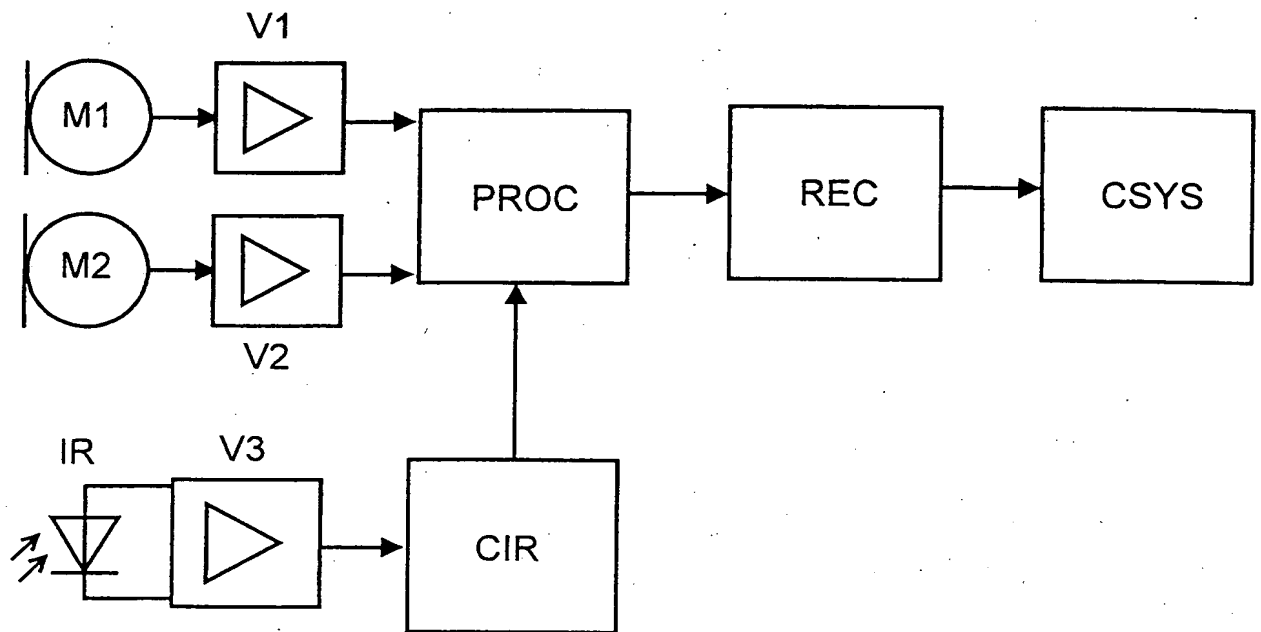


Fig. 1